

2. Sintaksna analiza - gramatike

Zadatak 1. (2.1.1 - deo)

Koja od sledećih sekvenci pripada jeziku opisanom zadatom gramatikom sa startnim neterminalom $\langle S \rangle$? Za svaki od slučajeva dati levo izvođenje, desno izvođenje i stablo izvođenja.

a) aacb

b) abcccb

1. $\langle S \rangle \rightarrow a \langle A \rangle c \langle B \rangle$

5. $\langle A \rangle \rightarrow \langle B \rangle a \langle B \rangle$

2. $\langle S \rangle \rightarrow \langle B \rangle d \langle S \rangle$

6. $\langle A \rangle \rightarrow a \langle B \rangle c$

3. $\langle B \rangle \rightarrow a \langle S \rangle c \langle A \rangle$

7. $\langle A \rangle \rightarrow a$

4. $\langle B \rangle \rightarrow c \langle A \rangle \langle B \rangle$

8. $\langle B \rangle \rightarrow b$

Rešenje

a) U prvom koraku izvođenja, u obzir za zamenu $\langle S \rangle$ dolaze:

- smena 1, čija desna strana počinje terminalom a, kao i
- smena 2, čija desna strana počinje neterminalom $\langle B \rangle$, koji se može kasnije zameniti primenom smene 3. čija desna strana takođe počinje terminalom a. (otpada zbog d)

$$\begin{array}{c} \langle S \rangle \Rightarrow a \langle A \rangle c \langle B \rangle \\ \uparrow \\ 1 \end{array}$$

$$\begin{array}{ccccc} \langle S \rangle \Rightarrow a \langle A \rangle c \langle B \rangle \Rightarrow a a c \langle B \rangle \Rightarrow a a c b \\ \uparrow \quad \quad \uparrow \quad \quad \uparrow \\ 1 \quad \quad 7 \quad \quad 8 \end{array}$$

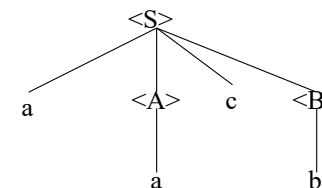
- Pošto je u svakom koraku izvođenja zamenjivan prvi neterminal s leve strane u svakoj sentencijalnoj formi, dobijeno izvođenje naziva se levim izvođenjem (engl. leftmost derivation), što se može posebno naznačiti na sledeći način:

$$\langle S \rangle \Rightarrow_{lm} a \langle A \rangle c \langle B \rangle \Rightarrow_{lm} a a c \langle B \rangle \Rightarrow_{lm} a a c b$$

- Ukoliko se u svakom koraku izvođenja zamenjuje krajnje desni neterminal u sentencijalnoj formi, dobija se desno izvođenje (engl. rightmost derivation):

$$\langle S \rangle \Rightarrow_{rm} a \langle A \rangle c \langle B \rangle \Rightarrow_{rm} a \langle A \rangle c b \Rightarrow_{rm} a a c b$$

- Stablo izvođenja je grafička predstava izvođenja. Izvođenje $\langle X \rangle \Rightarrow Y Z V$ se u stablu izvođenja predstavlja na taj način što čvor $\langle X \rangle$ ima čvorove naslednike Y, Z i V navedenim redosledom. U konkretnom slučaju, stablo izvođenja je:



- b) U prvom koraku mora se primeniti 1. smena (jer 2. smena sadrži terminal d koji se ne pojavljuje u ulaznoj sekvenci):

$$\begin{array}{c} \langle S \rangle \Rightarrow a \langle A \rangle c \langle B \rangle \\ \uparrow \\ 1 \end{array}$$

Varijante uparivanja c iz $a \langle A \rangle c \langle B \rangle$ sa c iz abcccb:

- Ukoliko se uparuje sa prvim c s leve strane u ulazu, treba da važi:

$$\langle A \rangle \xRightarrow{*} b \quad \text{što ne može}$$

- Ukoliko se uparuje sa drugim c s leve strane u ulazu, treba da važi:

$$\langle A \rangle \xRightarrow{*} bc \quad \text{što takođe ne može}$$

- Zaključujemo da ne važi $\langle S \rangle \xRightarrow{*} abcccb$, odnosno da pomenuta sekvenca ne pripada jeziku date gramatike.

Zadatak 3. (2.3.1)

Pronaći suvišne neterminale u sledećoj gramatici sa startnim simbolom $\langle S \rangle$.

1. $\langle S \rangle \rightarrow a\langle A \rangle\langle B \rangle\langle C \rangle$
2. $\langle S \rangle \rightarrow b\langle C \rangle\langle E \rangle\langle S \rangle$
3. $\langle S \rangle \rightarrow a\langle E \rangle$
4. $\langle A \rangle \rightarrow b\langle E \rangle$
5. $\langle A \rangle \rightarrow \langle S \rangle\langle C \rangle\langle D \rangle$
6. $\langle A \rangle \rightarrow d$
7. $\langle B \rangle \rightarrow d\langle F \rangle\langle S \rangle$
8. $\langle B \rangle \rightarrow a\langle B \rangle\langle C \rangle$
9. $\langle C \rangle \rightarrow a\langle E \rangle\langle S \rangle$
10. $\langle C \rangle \rightarrow b\langle E \rangle$
11. $\langle D \rangle \rightarrow a\langle A \rangle\langle C \rangle$
12. $\langle D \rangle \rightarrow d$
13. $\langle E \rangle \rightarrow a\langle C \rangle\langle E \rangle$
14. $\langle E \rangle \rightarrow \varepsilon$
15. $\langle F \rangle \rightarrow \langle A \rangle\langle B \rangle$
16. $\langle F \rangle \rightarrow a\langle F \rangle$

Analiza problema

Suvišni neterminali dele se na mrtve i nedostižne. Neterminal $\langle A \rangle$ je mrtav ako i samo ako se iz $\langle A \rangle$ ne može izvesti nijedna sekvenca terminalnih simbola, to jest:

$$\{w \mid \langle A \rangle \xRightarrow{*} w\} = \emptyset.$$

U suprotnom slučaju neterminal $\langle A \rangle$ je živ.

Neterminal $\langle A \rangle$ je nedostižan ako i samo ako se $\langle A \rangle$ ne pojavljuje ni u jednoj sentencijalnoj formi izvedenoj iz startnog neterminala $\langle S \rangle$, odnosno ako ne važi:

$$\langle S \rangle \xRightarrow{*} \alpha \langle A \rangle \beta$$

gde su α i β proizvoljne sekvence gramatičkih simbola dužine nula ili više. U suprotnom slučaju neterminal $\langle A \rangle$ je dostižan.

Suvišne neterminale i smene u kojima se oni pojavljuju moguće je ukloniti iz gramatike bez uticaja na jezik koji ona opisuje.

Rešenje

Prvo treba pronaći i ukloniti mrtve neterminale, pa onda nedostižne, jer uklanjanjem mrtvih neterminala neki od prethodno dostižnih neterminala mogu postati nedostižni.

Odredićemo sve žive neterminale date gramatike, preostali su mrtvi. Skup živih neterminala formiramo iterativno na sledeći način:

1. Inicijalno je skup živih neterminala \tilde{Z} prazan.
2. Razmatramo redom smene date gramatike: Ako je razmatrana smena oblika $\langle A \rangle \rightarrow \alpha$ pri čemu je desna strana prazna sekvenca ili sekvenca sastavljena isključivo od terminala i samo onih neterminala koji se pojavljuju u skupu \tilde{Z} (ako takvi postoje), dodajemo neterminal $\langle A \rangle$ u skup \tilde{Z} .
3. Postupak nastavljamo sve dok postoji promena u skupu \tilde{Z} , razmatrajući po potrebi svaku gramatičku smenu više puta.

4. Neterminali koji nisu živi pripadaju skupu mrtvih neterminala $M = V_N - \tilde{Z}$.

U zadatoj gramatici,

- $\langle A \rangle$, $\langle D \rangle$, $\langle E \rangle$ su sigurno živi jer se mogu zameniti terminalom
- $\langle S \rangle$ je živ na osnovu 3. smene
- $\langle C \rangle$ je živ na osnovu 9. smene

Mrtvi neterminali su dakle $\langle B \rangle$ i $\langle F \rangle$. Sve smene u kojima se pojavljuju $\langle B \rangle$ i $\langle F \rangle$ su nepotrebne pa iz gramatike uklanjamo smene 1, 7, 8, 15 i 16, pa ostaju smene:

2. $\langle S \rangle \rightarrow b\langle C \rangle\langle E \rangle\langle S \rangle$
3. $\langle S \rangle \rightarrow a\langle E \rangle$
4. $\langle A \rangle \rightarrow b\langle E \rangle$
5. $\langle A \rangle \rightarrow \langle S \rangle\langle C \rangle\langle D \rangle$
6. $\langle A \rangle \rightarrow d$
9. $\langle C \rangle \rightarrow a\langle E \rangle\langle S \rangle$
10. $\langle C \rangle \rightarrow b\langle E \rangle$
11. $\langle D \rangle \rightarrow a\langle A \rangle\langle C \rangle$
12. $\langle D \rangle \rightarrow d$
13. $\langle E \rangle \rightarrow a\langle C \rangle\langle E \rangle$
14. $\langle E \rangle \rightarrow \varepsilon$

Sada treba odrediti sve dostižne neterminale date gramatike, preostali su nedostižni. Skup dostižnih neterminala formiramo iterativno na sledeći način:

1. Inicijalno, skup dostižnih neterminala D sadrži samo startni neterminal.
2. Iz skupa D biramo neterminal $\langle A \rangle$ koji nije prethodno razmatran. Razmatramo sve smene sa levom stranom $\langle A \rangle$. U skup D dodajemo sve neterminale koji se pojavljuju na desnim stranama ovih smena.
3. Prethodni korak ponavljamo dok ne razmotrimo sve neterminale iz skupa D .
4. Neterminali koji nisu dostižni pripadaju skupu nedostižnih neterminala $N = V_N - D$.

U zadatoj gramatici:

- dostižan je $\langle S \rangle$
- iz $\langle S \rangle$ možemo izvesti sekvence sa $\langle C \rangle$ i $\langle E \rangle$ pa su i oni dostižni;
- iz smena za $\langle C \rangle$ i $\langle E \rangle$ ne možemo u skup D dodati nove neterminale; postupak se okončava.

Dakle nedostižni su $\langle A \rangle$ i $\langle D \rangle$. Iz gramatike uklanjamo sve smene u kojima se oni nalaze. Tako se gramatika svodi na:

2. $\langle S \rangle \rightarrow b\langle C \rangle\langle E \rangle\langle S \rangle$
3. $\langle S \rangle \rightarrow a\langle E \rangle$
9. $\langle C \rangle \rightarrow a\langle E \rangle\langle S \rangle$
10. $\langle C \rangle \rightarrow b\langle E \rangle$
13. $\langle E \rangle \rightarrow a\langle C \rangle\langle E \rangle$
14. $\langle E \rangle \rightarrow \varepsilon$

Zadatak 4. (2.4.1)

Naći desno-linearnu gramatiku bez suvišnih neterminala koja odgovara konačnom automatu sa slike:

	0	1	
→A	A, B	E	0
B	E	C	1
C	C	B, C	0
D	E	C, D	1
E	E	E	0

Analiza problema

Za zadati konačni automat moguće je odrediti odgovarajuću desno-linearnu gramatiku na sledeći način:

- Terminalnim simbolima odgovaraju ulazni simboli automata.
- Neterminalnim simbolima odgovaraju stanja automata. Startnom neterminalu odgovara startno stanje automata.
- Proizvoljnom prelazu automata iz stanja A u stanje B pod ulazom x, odgovara smena $\langle A \rangle \rightarrow x \langle B \rangle$.
- Proizvoljno stanje prihvatanja A uzrokuje dodavanje smene: $\langle A \rangle \rightarrow \varepsilon$.

Rešenje

Zadatom automatu odgovara sledeća gramatika:

- $\langle A \rangle \rightarrow 0 \langle A \rangle$
- $\langle A \rangle \rightarrow 0 \langle B \rangle$
- $\langle A \rangle \rightarrow 1 \langle E \rangle$
- $\langle B \rangle \rightarrow 0 \langle E \rangle$
- $\langle B \rangle \rightarrow 1 \langle C \rangle$
- $\langle B \rangle \rightarrow \varepsilon$
- $\langle C \rangle \rightarrow 0 \langle C \rangle$
- $\langle C \rangle \rightarrow 1 \langle B \rangle$
- $\langle C \rangle \rightarrow 1 \langle C \rangle$
- $\langle D \rangle \rightarrow 0 \langle E \rangle$
- $\langle D \rangle \rightarrow 1 \langle C \rangle$
- $\langle D \rangle \rightarrow 1 \langle D \rangle$
- $\langle D \rangle \rightarrow \varepsilon$
- $\langle E \rangle \rightarrow 0 \langle E \rangle$
- $\langle E \rangle \rightarrow 1 \langle E \rangle$

Lako ustanovljavamo da je neterminal $\langle E \rangle$, koji odgovara stanju greške mrtav, pa možemo ukloniti smene 3., 4., 10., 14. i 15. Neterminal $\langle D \rangle$ nije dostižan, pa možemo ukloniti i smene 11., 12. i 13. Konačni izgled gramatike je:

- $\langle A \rangle \rightarrow 0 \langle A \rangle$
- $\langle A \rangle \rightarrow 0 \langle B \rangle$
- $\langle A \rangle \rightarrow 1 \langle C \rangle$
- $\langle B \rangle \rightarrow 0 \langle E \rangle$
- $\langle B \rangle \rightarrow 1 \langle C \rangle$
- $\langle B \rangle \rightarrow \varepsilon$
- $\langle C \rangle \rightarrow 0 \langle C \rangle$
- $\langle C \rangle \rightarrow 1 \langle B \rangle$

- $\langle A \rangle \rightarrow 0 \langle B \rangle$
- $\langle B \rangle \rightarrow 1 \langle C \rangle$
- $\langle B \rangle \rightarrow \varepsilon$
- $\langle C \rangle \rightarrow 1 \langle B \rangle$
- $\langle C \rangle \rightarrow 1 \langle C \rangle$

Zadatak 5. (2.4.2)

Data je desno-linearna gramatika sa startnim simbolom $\langle A \rangle$:

- $\langle A \rangle \rightarrow ab \langle B \rangle$
- $\langle A \rangle \rightarrow ac$
- $\langle B \rangle \rightarrow \langle A \rangle$
- $\langle B \rangle \rightarrow \varepsilon$

- Transformisati datu gramatiku u regularnu gramatiku.
- Konstruisati nedeterministički konačni automat koji odgovara regularnoj gramatici.

Analiza problema

a)

Regularna gramatika koja opisuje isti jezik kao zadata desno-linearna gramatika može se dobiti transformacijom zadate gramatike po sledećim pravilima:

1.

$\langle A \rangle \rightarrow w$	zameniti sa	$\langle A \rangle \rightarrow w \langle \varepsilon \rangle$
		$\langle \varepsilon \rangle \rightarrow \varepsilon$
w – samo terminali		
$\langle \varepsilon \rangle$ – novouvedeni neterminalni simbol.		

2.

		$\langle A \rangle \rightarrow a_1 \langle a_2 \dots a_n B \rangle$
		$\langle a_2 \dots a_n B \rangle \rightarrow a_2 \langle a_3 \dots a_n B \rangle$
$\langle A \rangle \rightarrow a_1 a_2 \dots a_n \langle B \rangle$	zameniti sa	...
$n > 1$		$\langle a_{n-1} a_n B \rangle \rightarrow a_{n-1} \langle a_n B \rangle$
		$\langle a_n B \rangle \rightarrow a_n \langle B \rangle$
$\langle a_2 \dots a_n B \rangle, \langle a_3 \dots a_n B \rangle, \dots, \langle a_{n-1} a_n B \rangle, \langle a_n B \rangle$ – novouvedeni neterminali.		

3.

$\langle A \rangle \rightarrow \langle B \rangle$	
$\langle B \rangle \rightarrow \alpha_1$	$\langle A \rangle \rightarrow \alpha_1$
$\langle B \rangle \rightarrow \alpha_2$	zameniti sa $\langle A \rangle \rightarrow \alpha_2$
...	...
$\langle B \rangle \rightarrow \alpha_n$	$\langle A \rangle \rightarrow \alpha_n$
prethodno ukloniti smenu $\langle B \rangle \rightarrow \langle B \rangle$ ako eventualno postoji.	

Rešenje

Primenom pravila 1. na drugu smenu zadate gramatike, dobijamo:

1. $\langle A \rangle \rightarrow ab\langle B \rangle$
2. $\langle A \rangle \rightarrow ac\langle \varepsilon \rangle$
3. $\langle \varepsilon \rangle \rightarrow \varepsilon$
4. $\langle B \rangle \rightarrow \langle A \rangle$
5. $\langle B \rangle \rightarrow \varepsilon$

Primena pravila 2. na smene 1. i 2. dobijamo:

1. $\langle A \rangle \rightarrow a\langle bB \rangle$
2. $\langle bB \rangle \rightarrow b\langle B \rangle$
3. $\langle A \rangle \rightarrow a\langle C\varepsilon \rangle$
4. $\langle C\varepsilon \rangle \rightarrow c\langle \varepsilon \rangle$
5. $\langle \varepsilon \rangle \rightarrow \varepsilon$
6. $\langle B \rangle \rightarrow \langle A \rangle$
7. $\langle B \rangle \rightarrow \varepsilon$

Primena pravila 3. na smenu 6:

1. $\langle A \rangle \rightarrow a\langle bB \rangle$
2. $\langle bB \rangle \rightarrow b\langle B \rangle$
3. $\langle A \rangle \rightarrow a\langle C\varepsilon \rangle$
4. $\langle C\varepsilon \rangle \rightarrow c\langle \varepsilon \rangle$
5. $\langle \varepsilon \rangle \rightarrow \varepsilon$
6. $\langle B \rangle \rightarrow a\langle bB \rangle$
7. $\langle B \rangle \rightarrow a\langle C\varepsilon \rangle$
8. $\langle B \rangle \rightarrow \varepsilon$

b)

Analiza problema

Konačni, u opštem slučaju nedeterministički, automat koji prihvata jezik regularne gramatike dobijamo prema sledećim pravilima:

- Ulaznim simbolima automata odgovaraju terminalni gramatički simboli.
- Stanjima automata odgovaraju neterminali. Startnom stanju odgovara startni neterminal.

- Svakoј smeni oblika $\langle A \rangle \rightarrow a\langle B \rangle$ odgovara u automatu prelaz iz stanja A u stanje B za ulazni simbol a.
- Smena oblika $\langle A \rangle \rightarrow \varepsilon$ implicira da je A stanje prihvatanja automata.

Rešenje

	a	b	c	
$\rightarrow A$	bB, Cε			0
bB		B		0
B	bB, Cε			1
Cε			ε	0
ε				1